EFFECTIVE VOIGT MODEL ESTIMATION USING MULTIPLE RANDOM STARTING VALUES AND PARAMETER BOUNDS SETTINGS FOR *IN VIVO* HEPATIC ^{1}H MAGNETIC RESONANCE SPECTROSCOPIC DATA

H. Ratiney[1], A. Bucur[1], M. Sdika[2], O. Beuf[1], F. Pilleul[1], S. Cavassila[1]

[1]CREATIS-LRMN, CNRS UMR 5220, Inserm U630, Université de Lyon, Université Lyon 1, INSA-Lyon, Villeurbanne, France. [2]CRMBM CNRS UMR 6612, Marseille, France.

ABSTRACT

In vivo hepatic ^{1}H lineshapes modeled by the complex Voigt function are desirable to reduce systematic error and obtain accurate fits. However, the optimization procedure becomes challenging when the peak resonances overlap and the proportion of Gaussian to Lorentzian dampings is *a priori* unknown. In this context, nonlinear least-squares algorithms generally invoked in Magnetic Resonance Spectroscopy quantification are highly sensitive to the starting values and parameter bounds. To alleviate this sensitivity, multiple random starting values and parameter bounds settings are used to generate candidate solutions. The "best fit" fulfilling requirements on the cost function and damping factor final values is then selected among them. *Monte Carlo* studies and an *in vivo* hepatic ^{1}H signal quantification demonstrated the relevance of the proposed strategy.

Index Terms— Voigt Lineshape, optimization, starting values, model parameter bounds, Magnetic Resonance Spectroscopy

1. INTRODUCTION

Accurate quantification is a major step of *in vivo* proton NMR spectroscopic studies. A large number of the numerical modelling techniques fit the spectra/signals to assumed model functions [1-5]. The use of an inappropriate model leads to systematic errors in metabolite concentration estimates [6, 7]. Previous studies focused on *in vivo* experiments performed in the brain have stated that experimental lineshapes are better modelled by the Voigt functions [6]. In case of *in vivo* hepatic [1]H Magnetic Resonance Spectroscopic (MRS) data, the MR spectrum exhibits broad patterns which also induce the use of a Voigt lineshape model function for its quantification. While the Voigt lineshape function is a more suited and flexible model, it also introduces more correlation terms between the model parameters. Moreover when dealing with Voigt lineshapes, the proportion of Gaussian to Lorentzian is *a priori* unknown. These sources of uncertainty as well as the parameter interdependences make the choice of the starting values and priori knowledge crucial in order to avoid local minima. The time-domain algorithms available in the literature for MRS quantification have proposed several approaches to set-up the starting values. In AMARES [2], starting values of the nonlinear parameters are obtained by peak-picking on the spectrum display so that they are close to the solution. In QUEST [3] or AQSES [4], the use of a metabolite basis-set alleviates somehow user's involvements. Indeed, the metabolite basis-set constitutes both the prior knowledge included in the model function and the starting values of the nonlinear parameters. In QUEST, the upper and lower bounds as well as extra-damping starting values sometimes need to be tuned to better adapt the basis-set to the *in vivo* signal. In the AQSES framework which uses an improved implementation of VARPRO [5], it was found that good initial values for the non linear parameters were zeros. Altogether, in all these different algorithms, only a single set of starting values is used.

The goal of this paper was to demonstrate the relevance of an effective strategy to set up starting values and parameter bounds when fitting, via a minimization algorithm of the Levenberg-Marquardt type, a mixture of Voigt functions as the one encountered in hepatic MR spectroscopy. A fully automatic fitting scheme based on the Voigt model function which is robust regarding the influences of the starting values and the model parameter bounds is derived. Finally, quantification of short-echo time ^{1}H MRS human liver signals acquired at 1.5T is demonstrated.

2. METHOD

2.1. The core algorithm

The quantification procedure is based on a nonlinear least-squares algorithm (Levenberg-Marquardt, MATLAB 7.4) that fits the *in vivo* time-domain signal to a combination of K Voigt lineshape functions (K corresponds to the number of resonances in the *in vivo* spectrum). Each Voigt component is characterized by the combination of Lorentzian and Gaussian functions.

$$\hat{y}_n = \exp(i\phi_0) \sum_{k=1}^{K} a_k \exp(\alpha_k t_n + \beta_k^2 t^2{}_n + i2\pi f_k t_n) ,$$ Eq. 1

$n = 0 \dots N\text{-}1$

The model parameters a_k, α_k, β_k, f_k, ϕ_0 correspond to the amplitudes, the Lorentzian and Gaussian damping factors, the frequencies and the zero-order phase. N is the number of data-points. $t_n = nt_s + t_0$ are the sampling times, in which t_0 is the dead-time of the receiver and t_s the sampling interval and $i^2 = -1$. The optimization procedure aims to obtain the model parameters that minimize the cost function:

$$\chi^2_{UP, StartL, StartG} = \sum_{N=0}^{N-1} (y_n - \hat{y}_N)^2 ,$$ Eq. 2

where y_n are the measured complex-value time-domain data-points, $\hat{y}_n$ are the complex-value data-points of the sum of the Voigt functions. The cost function is indexed by 4 sets of values: the upper bounds UP of the Gaussian and Lorentzian damping parameters, and their starting values respectively $Start_L$ and $Start_G$. Frequencies need first to be provided by the user. Then, the global damping parameters are guessed from the Full Width at Half Maximum (FWHM) measurements made on the peaks corresponding to the user–defined set of frequencies. The global damping parameter of each component is roughly estimated as the average of theoretical Gaussian and Lorentzian damping factors according to the following approximation:

$$\gamma_{global,k} = \pi \frac{FWHM}{2} \left(1 + \frac{1}{2\sqrt{\ln 2}}\right) \approx \alpha_k + \beta_k$$ Eq. 3

But, we emphasis once again that the relative proportions of Lorentzian and Gaussian damping factors are *a priori* unknown in the global damping factor. After setting arbitrarily or randomly these proportions, starting values for the linear parameters (amplitudes and phases) can be deduced by solving a linear least-squares (LS) problem as used in [3,4]. Frequency parameters were constrained within an interval of ± 5Hz around the frequency starting values.

2.2. Influence of the starting values and upper bounds of the Voigt model parameter on estimates for a noiseless signal

A signal mimicking an *in vivo* [1]H hepatic signal acquired at 1.5T was generated for the *Monte Carlo* study. It consists of a weighted sum of ten Voigt functions. The model parameters of the Voigt functions are displayed in Table 1. The zero-order phase was set to zero. This signal was quantified two hundred times, with varying sets of frequency starting values. These sets were randomly chosen within an interval of ± 2Hz around the true frequencies using a uniform distribution. Each time, the procedure defined in 2.1 was used with a single set of starting values and bounds. The starting values $Start_L$ and $Start_G$ were arbitrarily chosen in equal proportion in the measured global damping factors.

The upper bounds for all the damping factors were set to 130Hz. These upper bounds were chosen empirically to match the observed hepatic liver signal properties. The sensitivity of the optimization problem to the starting values was pointed out by drawing the distributions of the cost functions, of the amplitude parameter estimates and of the damping factor estimates for the two hundred fits of the noiseless signal.

	f (ppm)	α (Hz)	β (Hz)	a (a.u.)
1	5.07	2.96	107.99	248
2	4.68	0.36	41.29	873
3	4.58	18.89	10.17	388
4	3.84	0.56	37.81	78
5	3.62	0.02	53.42	93
6	3.17	0.00	30.43	90
7	2.09	0.00	43.19	80
8	1.72	0.02	79.63	70
9	1.21	0.01	38.88	694
10	0.82	0.01	48.95	211

Table 1: Model parameters of the ten Voigt functions in the simulated signal considered for the *Monte Carlo* study.

2.3. Strategy to handle the starting values and upper bounds effects on the Voigt model parameter estimates

Two strategies to handle the starting values and upper bounds effects on the Voigt model parameter estimates were designed and compared.
- *Single user-defined starting values and bounds (SSV).*
- *Multiple randomly chosen starting values (MSV) in a suited range and selection of the "best fit".*

2.1.1. SSV
The first strategy called *SSV* corresponds to the usual set-up of starting values. Well designed starting values for the frequencies were obtained by peak-picking the frequencies on the spectrum display. The upper bounds of the damping factor parameters were fixed to 130 Hz (value empirically chosen to match the hepatic liver signal properties).

2.1.2. MSV
In order to reduce the risk of convergence to local minima, the fittings were repeated automatically with several combinations for the starting values and upper bounds. A total of about seven-hundred combinations were randomly set-up. Since no prior knowledge is available, uniform distributions were used to draw these combinations according the following framework.
- For each component, one hundred global damping factors $\gamma_{global,k}$ were drawn using a uniform distribution within an interval of ± 30% around the values estimated according to 2.1 to take into account uncertainties on the

estimation of $\gamma_{\text{global},k}$. Then the starting values $Start_L$ and $Start_G$ were set-up in random proportion within the global damping factor, according to Eq.4.

$$\begin{cases} \alpha_k = \lambda\gamma_{\text{global},k} \\ \beta_k = (1-\lambda)\gamma_{\text{global},k} \end{cases} \quad \text{Eq. 4}$$

where λ follows a uniform distribution over [0,1].

- Several upper bounds for the Lorentzian and Gaussian damping factor parameters (UP_L and UP_G) were set-up and tested. The first sets were set to the starting values enlarged by 20Hz. They were then independently incremented by step of 20Hz till the minimum between twice the starting value and 130Hz was reached.

At the end, a unique set of model-parameter estimates was selected among solutions obtained by the different UP, $Start_L$, $Start_G$ combinations. This set of model parameters, corresponding to the "best fit", fulfil the requirement of producing the lowest cost function and its damping factors being lower than 99% of their respective upper bounds UP.

2.4. *Monte Carlo* simulations

The comparison of both strategies was performed with the aid of *Monte Carlo* simulations. A total of fifty realizations of a white Gaussian distributed noise were added to the simulated signal presented in Table 1. The noise level was chosen corresponding to signal-to-noise ratios (SNRs) of 353:1 compared to the total amplitude. This SNR corresponds approximately to the *in vivo* measured SNRs. The Lorentzian and the Gaussian damping factors, the frequencies, the amplitudes and the zero-order phase are the parameter to estimate. The statistical performances of the two strategies were evaluated and compared through the computation of the means, the biases and the standard deviations of all the estimates.

2.5. Application to *in vivo* hepatic spectroscopic signals

The MR spectroscopy signal used to demonstrate the effects of the proposed strategies was derived from an acquisition on a patient with suspected diffuse liver diseases [8]. The MR experiments were performed on a clinical 1.5 T Symphony system (Siemens Medical Solutions, Erlangen, Germany) using phased-array body coils. Localized MRS acquisitions were performed using a short–echo time respiration trigged PRESS sequence (TR/TE 1500/30 ms, about 6 minutes acquisition time, depending on the patient's respiratory cycle). The Eddy current effects were corrected using the water signal as reference. Ten components were selected to fit the saturated lipids (0.9ppm, 1.3ppm and 1.8ppm), the unsaturated lipids (2.1ppm, 5.2ppm), the water residue (2 components around 4.7ppm) and the metabolite contributions, See Figure 3. The assignations of the resonances referred to the published values in [9, 10].

3. RESULTS

3.1. Study on the noiseless signal

The distribution of the cost functions related to the 200 fits of the simulated noiseless signal is displayed on the left column of Fig 1. The distributions of the amplitude parameter estimates versus the frequency parameter estimates (middle column) and the damping factor estimates (right column) for the resonances 1, 5 and 10 of Table 1 were also reported. The fitting of this noiseless signal using a complex Voigt lineshape model corresponds to a non-convex optimization problem, with multiple feasible regions and multiple locally optimal points within each region. This is demonstrated by the distributions shown in Figure 1. Note that these distributions are related to the properties of the resonance peak itself and its interactions with the other peaks. These interactions can be partially characterized by the correlation term between the model function parameters computed from the Fisher Information matrix (results not shown).
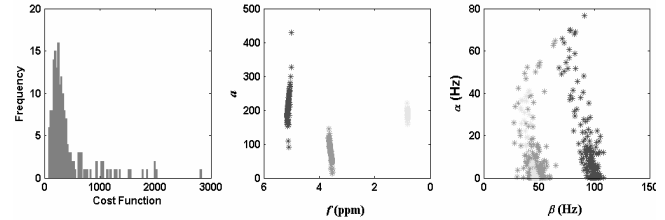

Figure 1: Distributions related to the two-hundred solutions of the repeated fits of the noiseless signal mimicking an *in vivo* [1]H hepatic signal acquired at 1.5T of a) the cost functions (histogram), b) the amplitude and frequency estimates and c) the Lorentzian and Gaussian damping factor estimates, for resonances 1,5,10 of Tab 1.

3.2. *Monte Carlo* simulations on noisy signals

Means and standard deviations for the amplitude, the Lorentzian and Gaussian damping factor parameter estimates are displayed for peak n°1, 5 10 on Figure 2. The bias is the distance between the mean value and the solid line corresponding to the actual values. The biases were noticeably reduced using the MSV strategy, reflecting the robustness of this approach. The *Monte Carlo* showed that, in general, the standard deviations were also reduced.

3.3. Application to *in vivo* hepatic spectroscopic signals

The MSV strategy was applied on an *in vivo* liver spectroscopy signal. The final cost functions for all the tested starting values and bounds are displayed in Figure 3. The "best fit" was highlighted among all the candidate solutions and the corresponding signal estimate is displayed. The main groups of resonances originating from different types of lipids, compounds of choline, TMAO, glucose and glycogen were successfully quantified.
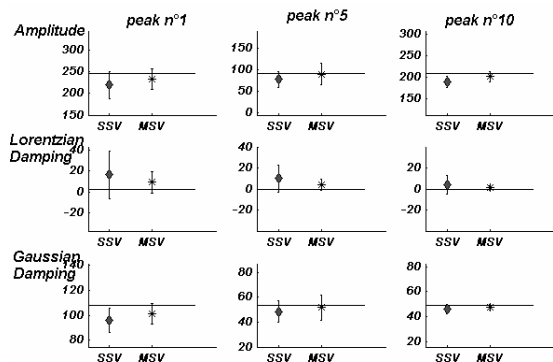
Figure 3: *Monte Carlo* simulation results (mean and standard deviation of 50 noisy signals) comparing the use of a single set of starting values and bounds (*SSV*) to the use of multiple starting values and bounds and "best fit" selection (*MSV)* for peak n°1, 5 10. Solid lines design the actual values.

4. CONCLUSION

An effective framework using multiple random starting values and bounds was proposed for Voigt model estimation in hepatic ^{1}H magnetic resonance spectroscopic signals.

Indeed *Monte Carlo* studies showed substantial bias reductions and general standard deviation decreases for the amplitude, Lorentzian and Gaussian damping factor estimates. Consequently, a MSV type approach is highly advisable when dealing with Voigt lineshapes. The counterpart of this approach is its computational time.

However, parallelization of the procedure can attenuate this drawback and is straightforward to implement as the starting values and bounds settings are independent of each other. Finally, this fully automated quantification technique will enable robust analyses of large amount of data required by clinical studies or MR spectroscopic imaging studies.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] S.W. Provencher, "Estimation of metabolite concentrations from localized in vivo proton NMR spectra", *Magn. Reson. Med.*, 30(6), pp 672-679, 1993.

[2] L. Vanhamme, A. van den Boogaart, and S. Van Huffel," Improved method for accurate and efficient quantification of MRS data with use of prior knowledge", *J Magn Reson*. 129(1), pp 35-43, 1997.

[3] H. Ratiney, M. Sdika, Y. Coenradie, S. Cavassila, D. van Ormondt, and D. Graveron-Demilly, " Time-domain semi-parametric estimation based on a metabolite basis set." *NMR Biomed*. 18, pp 1-13, 2005.

[4] J.B. Poullet, D.M. Sima, A.W. Simonetti, B. De Neuter, L. Vanhamme, P. Lemmerling and S. Van Huffel, "An automated quantitation of short echo time MRS spectra in an open source software environment: AQSES.*", NMR Biomed*., 20, pp 493-504, 2007.

[5] J.W van der Veen, R. de Beer, P.R. Luyten and D. van Ormondt," Accurate quantification of in vivo 31P NMR signals using the variable projection method and prior knowledge", *Magn. Reson. Med.*, 6, pp 92-98, 1988.

[6] I. Marshall, S.D. Bruce, J. Higinbotham, A. MacLullich, J.M. Wardlaw, K.J. Ferguson and J. Seckl, "Choice of spectroscopic lineshape model affects metabolite peak areas and area ratios", *Magn. Reson. Med.*, 44, pp 646-9, 2000.

[7] C. Cudalbu, S. Cavassila, H. Rabeson, D. van Ormondt and D. Graveron-Demilly, "Influence of measured and simulated basis sets on metabolite concentration estimates."*NMR Biomed*, DOI:10.1002/nbm.1234, 2007.

[8] F. Pilleul, G. Chave, J. Dumortier, J.Y. Scoazec, P.J. Valette, "Fatty infiltration of the liver. Detection and grading using dual T1 gradient echo sequences on clinical MR system.", *Gastroenterol Clin Biol*., 29, pp 1143-1147, 2005.

[9] B. Martinez-Granados, D. Monleón, M.C. Martínez-Bisbal MC, J.M. Rodrigo, J. del Olmo, P. Lluch, A. Ferrández, L. Martí-Bonmatí, B. Celda. "Metabolite identification in human liver needle biopsies by high-resolution magic angle spinning 1H NMR spectroscopy.", *NMR Biomed* , 19, pp90-100, 2006.

[10] I.J. Cox, A. Sharif, J.F. Cobbold, H.C. Thomas and S.D. Taylor-Robinson*,* "Current and future applications of in vitro magnetic resonance spectroscopy in hepatobiliary disease." *World J. Gastroenterol.*, 12(30), pp 4773-4783, 2006.
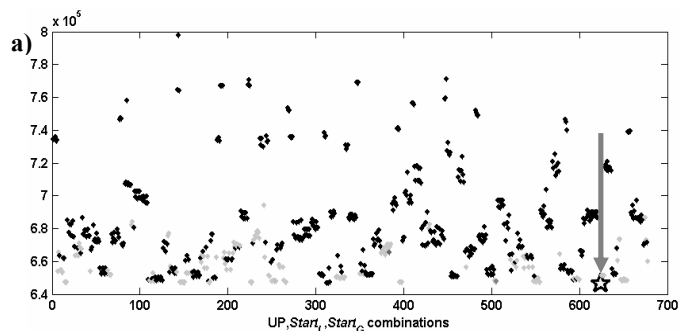
Figure 3: Quantification of a ^{1}H water-suppressed spectrum acquired at 1.5T from the right hepatic lobe of a patient a) Cost function distribution of all the candidate solutions, in gray the cost function for which the damping parameters were lower than the upper bounds, in black the one for which they reached the upper bound, the star corresponds to the selected "best fit" displayed in b) as a sum of Voigt lineshape resonances (dotted line), the original spectrum (dark gray), individual components (light gray) and the residue (black bold line).